Hammerstein-Wiener System Estimator Initialization

Ph. Crama, J. Schoukens
VUB, Fakulteit Toegepaste Wetenschappen, Vakgroep ELEC
Pleinlaan 2, B-1050, Brussel, Belgium
e-mail: Philippe.Crama@vub.ac.be

Abstract

In nonlinear system identification, the system is often represented as a series of blocks linked together. Such block-oriented models are built with static nonlinear subsystems and linear dynamic systems. This paper deals with the identification of the Hammerstein-Wiener model, which is a block-oriented model where a linear dynamic system is surrounded by two static nonlinearities at its input and output. The proposed identification scheme is iterative and will be demonstrated on measurements.

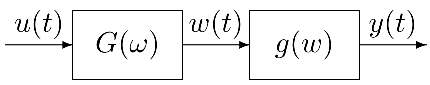

Figure 1: A Wiener system

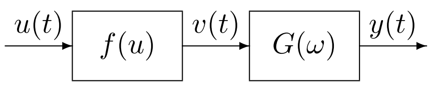

Figure 2: A Hammerstein system

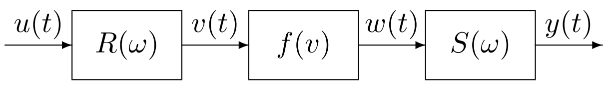

Figure 3: A Wiener-Hammerstein system

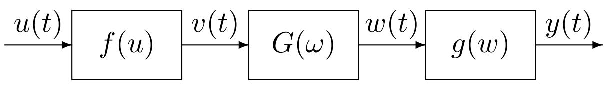

Figure 4: A Hammerstein-Wiener system

1 Introduction

In modern applications, devices and systems are increasingly used in regions where they can't be accurately described with a linear model anymore. This drives a growing need for models and modeling techniques able to adequately describe these system's behavior. Estimating nonlinear systems is a very broad problem : it is impossible to propose a structure able to describe efficiently every possible system. Hence, the scope is often reduced to focus on nonlinear model structures built with static nonlinearities and linear dynamic systems only.

The simplest examples of such systems are Wiener systems (Figure 1) and Hammerstein systems (Figure 2). When three blocks are combined, two different structures can be defined : the Wiener-Hammerstein system (Figure 3) and the Hammerstein-Wiener system (Figure 4). Hammerstein-Wiener systems describe linear systems surrounded by an input and an output static nonlinearity.

These block-oriented nonlinear models have been studied for some time now [1, 2, 3, 4, 5, 6], but there isn't much literature about Hammerstein-Wiener model identification [7, 8, 9], though this model is sometimes used for model-based predictive control [10].

This paper describes an iterative algorithm for the identification of Hammerstein-Wiener systems as defined in Figure 4. Even though the first step of the iteration can be proven to be biased, it appears that the method converges to the correct solution when the model structure includes the true system and when there is no noise. Another empirical constatation made on measurements is that the proposed method also allows to distinguish between Wiener (Figure 1) and Hammerstein (Figure 2) systems, thus alleviating the model structure selection problem. Since this method does not globally optimize the parameters of the model to minimize the cost function, its parameters should be used as initialization for a more complex estimator optimizing all parameters at once.

In the next section, some theoretical concepts will be introduced along with the reason why the first iteration of the algorithm produces a biased result in certain cases. Section 3 presents the algorithm used to identify the Hammerstein-Wiener system. In Section 4, the algorithm is applied to measurements and Section 5 demonstrates the ability of the algorithm to distinguish a Wiener system from a Hammerstein system.

2 Theoretical Background

This section introduces some necessary theory to explain some features of the presented algorithm. This topic is studied in much more detail in [11]. As any nonlinear system identification method, the identified model is optimized for a certain class of input signals. The signals used here are random phase multisines which can be tuned to the spectrum and to the amplitude distribution necessary for the considered application [12, 13, 14].

2.1 Random Phase Multisines

The proposed excitation signal, a random phase multisine, is defined in (1). The user can choose the amplitude distribution and basic frequency, but the phases are random.

$$u(t) = \sum_{k=-N}^{N} U_k \mathrm{e}^{\mathrm{j}\left(2\pi \frac{f_{\max}}{N} kt + \phi_k\right)} \tag{1}$$

In (1), U_k are the user-defined amplitudes with $U_k = U_{-k}$ and $\phi_k = -\phi_{-k}$ are the random phases such that $E\left[\mathrm{e}^{\mathrm{j}\phi_k}\right] = 0$.

2.2 The Related Linear Dynamic System

Working with random phase multisines allows to classify contributions of nonlinearities into “bias contributions” and “stochastic contributions” in (2) [11, 15]. This in turn gives rise to the definition of the Related Linear Dynamic System G_{R}. This is an important result : the Related Linear Dynamic System is easily computed and gives an approximation of the true linear system G inside the Hammerstein-Wiener system. It will be explained further on under which special conditions this approximation can be equal to the true value of G.

$$\begin{aligned} \frac{Y(\omega)}{U(\omega)} &= G_0(\omega) + G_{\mathrm{B}}(\omega) + G_{\mathrm{S}}(\omega) \\ &= G_{\mathrm{R}}(\omega) + G_{\mathrm{S}}(\omega) \end{aligned} \tag{2}$$

In (2), $G_0(\omega)$ is the underlying linear system (equal to $\lambda_1\mu_1 G(\omega)$ for the Hammerstein-Wiener system defined in Figure 4 and (3)), $G_{\mathrm{B}}(\omega)$ are the bias contributions, $G_{\mathrm{S}}(\omega)$ are the stochastic contributions and $G_{\mathrm{R}}(\omega)$ is the RLDS.

2.2.1 The RLDS of a Hammerstein-Wiener System

Because the RLDS of the system will be used as approximation of the linear system $G(\omega)$ inside the Hammerstein-Wiener system, its expected value should be studied. Here, the third Volterra kernel [16] of a Hammerstein-Wiener system will be computed. It will be shown that contrarily to the Wiener-Hammerstein case (see Figure 3 [11, 14, 15]), the third Volterra kernel can introduce a bias contribution into the RLDS that will make it different from the underlying linear system.

Consider $f(u)$, $g(w)$ defined by (3) : then the third Volterra kernel $H_3(\omega_1, \omega_2, \omega_3)$ is given by (4).

$$\left\{ \begin{aligned} f(u) &= \sum_{n=1}^{q_f} \lambda_n u^n \\ g(w) &= \sum_{n=1}^{q_g} \mu_n w^n \end{aligned} \right. \tag{3}$$

$$\begin{aligned} H_3(\omega_1, \omega_2, \omega_3) &= \mu_1 G(\omega_1 + \omega_2 + \omega_3)\lambda_3 + \\ &2\mu_2 G(\omega_1 + \omega_2) G(\omega_3) \lambda_2 \lambda_1 + \\ &\mu_3 G(\omega_1) G(\omega_2) G(\omega_3) \lambda_1^3 \end{aligned} \tag{4}$$

Putting $\omega = \omega_1 + \omega_2 + \omega_3$, we get that

$$\begin{aligned} &Y_3(\omega) = \\ &\quad \sum_{\omega_1}\sum_{\omega_2} \mu_1 G(\omega)\lambda_3 \\ &\quad\quad U(\omega - \omega_1 - \omega_2) U(\omega_1) U(\omega_2) + \\ &\quad \sum_{\omega_1}\sum_{\omega_2} \mu_2 G(\omega - \omega_1) G(\omega_1) \lambda_2 \lambda_1 \\ &\quad\quad U(\omega - \omega_1 - \omega_2) U(\omega_1) U(\omega_2) + \\ &\quad \sum_{\omega_1}\sum_{\omega_2} \mu_3 G(\omega - \omega_1 - \omega_2) G(\omega_1) G(\omega_2) \lambda_1^3 \\ &\quad\quad U(\omega - \omega_1 - \omega_2) U(\omega_1) U(\omega_2) \end{aligned} \tag{5}$$

This Volterra kernel contributes to the output as shown in (5). As explained in [11, 15], this contribution contains a bias term $G_{\mathrm{B}}^3(\omega)U(\omega)$ which appears when $\omega_1 = -\omega_2$ (6). This equation proves that if both static nonlinearities contain an even term, the bias contributions will not be proportional to the underlying linear system any more.

$$\begin{aligned}
G_{\mathrm{B}}^{3}(\omega) = {} & \mu_1 G(\omega) \lambda_3 \sum_{\omega_1} |U(\omega_1)|^2 + \\
& \mu_2 \lambda_2 \lambda_1 \sum_{\omega_1} G(\omega - \omega_1) G(\omega_1) |U(\omega_1)|^2 + \\
& \mu_3 G(\omega) \lambda_1^3 \sum_{\omega_1} |U(\omega_1)|^2 \\
= {} & C\left(\mu_1, \lambda_1, \mu_3, \lambda_3, U(\omega)\right) G(\omega) + \\
& \mu_2 \lambda_2 \lambda_1 \sum_{\omega_1} G(\omega - \omega_1) G(\omega_1) |U(\omega_1)|^2 \qquad (6)
\end{aligned}$$

In practice, the introduced error seems to be small enough and doesn't disturb the algorithm too much. Once the static nonlinearities are estimated (Section 3.2.2), this knowledge is used to lower the unwanted bias contributions. If one of the static nonlinearities is odd (e.g. $\lambda_{2n} = 0$), the unwanted term in (6) disappears, and the RLDS is proportional to the underlying linear system.

3 The Estimation Procedure

The estimation of the system parameters is an iterative process, where each iteration consists of two steps : identifying a parametric model for a linear system and solving a linear least squares problem to identify the static nonlinearities. The estimation of the linear system uses the estimation of the static nonlinearities of the previous iteration, while the estimation of the parameters of the static nonlinearities depends on the parameters of the linear system identified in the same iteration.

3.1 Parameterization of the Static Nonlinearities

The static nonlinearity $v = f(u)$ is parameterized using basis functions that can be freely chosen in (7) by the user as long as $f(u)$ is linear in the parameters. Possible choices of $f_l(u)$ include u^l or triangle functions (illustrated in Figure 5). With these triangle functions, it is possible to obtain a piecewise linear approximation of the nonlinear characteristic. However, only the height of the triangles are free, the centers have to be chosen a priori, otherwise the problem would be nonlinear in the parameters.

$$v(t) = \sum_{l=1}^{q_f} f_l(u(t)) \qquad (7)$$

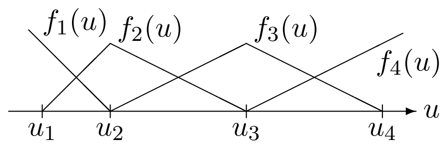

Figure 5: Triangle functions usable as basis functions for the static nonlinearities, yielding a piecewise linear approximation of the nonlinear characteristic

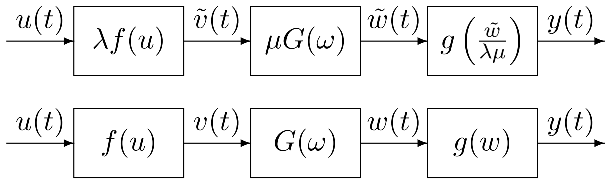

Figure 6: Two equivalent Hammerstein-Wiener systems

The static nonlinearity $g(w)$ is not identified directly because this would cause the estimation process to become nonlinear in the parameters. Rather than facing a nonlinear estimation problem for an algorithm that only proposes starting values for a nonlinear estimator, the problem is kept linear in the parameters by identifying the inverse $w = g^{-1}(y)$. The inverse of the output nonlinearity is parameterized using basis functions (which may be different from the ones used for the input nonlinearity $v = f(u)$) in (8). It follows from this parameterization that implicitly, the estimation algorithm assumes that the output static nonlinearity is invertible.

$$w(t) = \sum_{l=1}^{q_g} g_l(y(t)) \qquad (8)$$

3.2 The Algorithm

When estimating a Hammerstein-Wiener model (or any block structured model with internal signals that cannot be measured), there can be several degenerations. These are due to the fact that the gain of the system can be arbitrarily distributed over the different building blocks : a Hammerstein-Wiener system with $\tilde{f}(u) = \lambda f(u)$, $\tilde{G}(\omega) = \mu G(\omega)$ and $\tilde{g}(w) = g(\frac{w}{\lambda\mu})$ is the same as a Hammerstein-Wiener system with $f(u)$ as first nonlinearity, $G(\omega)$ as linear dynamic system and $g(w)$ as output nonlinearity (Figure 6). This shows that input-output measurements allow to reconstruct the different blocks' characteristic only up to an unknown scaling constant.

3.2.1 Estimating the Linear Part

At iteration k, the estimates of the static nonlinearities of iteration $k-1$ are used to guess the signals $v(t)$ and $w(t)$. These two signals are then used to identify a linear model that explains $w(t)$ as output for an input of $v(t)$. This identification can be done with any estimation method : in this case, frequency domain identification methods were used to take advantage of the periodic character of the excitation.

This parametric estimation step is very important : because of errors in the estimate of the static nonlinearities, the estimates of $v(t)$ and $w(t)$ are disturbed by a small bias due to the errors in the estimates of the nonlinear parts and by stochastic nonlinear contributions that behave like noise. If a non parametric model of the linear system is built by dividing the Fourier coefficients of the output by the Fourier coefficients of the input, the algorithm converges immediately to a biased estimate of the Hammerstein-Wiener system. The smoothing effect of the parametric estimation of the linear system is responsible for the convergence of the method : the effect of the stochastic contributions is removed (since the estimation of the linear system uses a consistent estimator [14]) so that only the bias contribution remains. This contribution is then lowered in the next step.

The iterative algorithm is started with $G^{[0]}(\omega)$ defined in (9). As explained in Section 2.2.1, this starting value is biased in certain cases. However, this fact has never prevented the algorithm to converge to the correct values in noiseless simulations on systems without model error and to good results on noisy measurements.

$$\begin{aligned} G^{[0]}(\omega) &= G_\theta\left(\omega;\theta^{[0]}\right) \\ \theta^{[0]} &= \arg\min_\theta \left|G_\theta(\omega;\theta) - G^{[0]}_{\text{NP}}(\omega)\right|^2 \\ G^{[0]}_{\text{NP}} &= \frac{Y(\omega)}{U(\omega)} \end{aligned} \tag{9}$$

3.2.2 Estimation of the Static Nonlinearities

Once the first step of iteration k is completed, the estimate of the static nonlinearities can be updated. This is done with a method similar to the one used in [17] for Hammerstein system identification.

The parameterization of the static nonlinearities (7), (8) has been chosen to be able to express an estimation problem that is linear in the parameters. The underlying idea is to compute the intermediate signal $w(t)$ in two different ways. The first possibility computes $w_u(t)$ as a linear combination of the input filtered by the fictional Hammerstein systems $G^{[k]}[f_l(u(t))]$ (10). The second possibility consists of writing $w_y(t)$ as a linear combination of the output filtered by the basis functions of the inverse of the output nonlinearity (11). Both (10) and (11) have been written in the frequency domain, once more to take advantage of the periodic character of the input signal : $w_u(t)$ can only represent $w(t)$ with accuracy at the frequencies where the model of the linear system was estimated, hence a filtering operation is needed. This filtering is (in the frequency domain for periodic excitations) as simple as selecting the correct Fourier coefficients.

$$W_u(\omega) = \sum_{l=1}^{q_f} a_l G^{[k]}(\omega)\mathcal{F}\{f_l(u(t))\} \tag{10}$$

$$W_y(\omega) = \sum_{l=1}^{q_g} b_l \mathcal{F}\{g_l(y(t))\} \tag{11}$$

$$e(\omega) = W_y(\omega) - W_u(\omega) \tag{12}$$

$$V(a_1,\ldots,a_{q_f},b_1,\ldots,b_{q_g}) = \sum_{l=1}^{N} |e(\omega_l)|^2 \tag{13}$$

These two expressions of $W(\omega)$ ((10) and (11)) have to be equal, hence the difference (12) has to be minimal, allowing to estimate the parameters of the static nonlinearities to ensure this equality in least-squares sense. The cost function that is minimized to solve this problem is given by (13). This minimization has to be done with an additional constraint to avoid the trivial solution $a_l = 0, b_l = 0$. This is of no consequence on the quality of the final solution because anyway the parameters may be freely scaled due to the gain indetermination mentioned at the beginning of this section.

4 The Algorithm applied to Measurements

This section presents the measurement setup used to test the proposed method on real-world data.

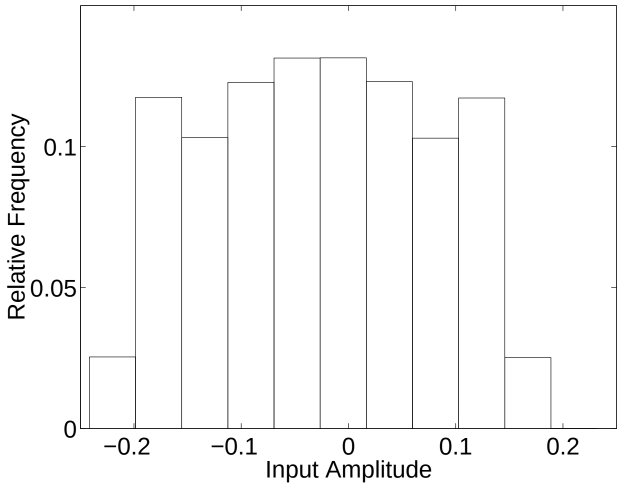

Figure 7: Amplitude distribution in 11 bins of the input signal

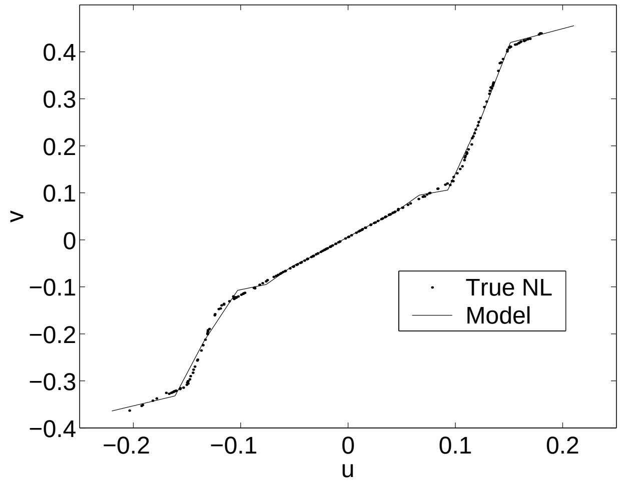

Figure 8: Identification result for the input nonlinearity $f(u)$ of the measured Hammerstein-Wiener system

4.1 The Measurement Setup

A laboratory setup has been assembled, consisting of two custom-made static nonlinearities and a Brüel & Kjær Octave Filter set with a center frequency of 125Hz. The highest frequency component of the input spectrum in the experiments has been kept below 200Hz, to make sure that the nonlinearities would indeed be static.

The measurements were controlled by a computer using an arbitrary waveform generator (HP E1445A) and sampled with two ADC cards (HP E1437A). The sampling frequency was 9.765625kHz, well above the Nyquist frequency for the used excitation. One period of 65536 points of four different signals was measured.

Figure 7 shows that the input multisines were given a special amplitude distribution to get a better estimate of the input nonlinearity for high input amplitudes : the phase values were optimized to get more points of the input with values at the extreme values of the distribution.

4.2 Results

For each of the four input signals, a first estimation of the measured Hammerstein-Wiener system was made. Both static nonlinearities were modeled using a piecewise linear function. The breakpoints ($u_1 \ldots u_4$ in Figure 5) were chosen automatically for equal support [1]. The interval between the minimum and the maximum value of the input of $f(u)$ was divided in bins in such a way that in each bin there would be the same number of input samples. This ensures that each parameter of the static nonlinearities is supported by the same number of measurements, which avoids numerical problems.

Figure 8, Figure 9 and Figure 10 show the results of the algorithm on the measurement data. On each of these plots, the full line represents the identified model and the dots represent the true values : on our laboratory setup, we were able to measure the normally unavailable signals $v(t)$ and $w(t)$. Obviously, these were used for comparison purposes only, but not for the identification. The gain degeneration mentioned in Section 3 was resolved by scaling the model in each plot in order to make the comparison between the true system and the model the easiest possible. These figures all show that the result is very good : the models all predict the true systems very well. The biggest imperfections in Figure 8 come from the non ideal breakpoint selection : the true system is a piecewise linear function, but the breakpoints of the model do not fall on the breakpoints of the true system. This causes the model to cut the corners of the characteristic, as can be seen most clearly in Figure 8 for input values around -0.1.

5 Identifying a Wiener model

The proposed algorithm was also applied on measurements done on a Wiener system as defined in Figure 1 [17]. The results of this identification procedure is shown in Figure 11, Figure 12 and Figure 13.

A Wiener system is a special case of a

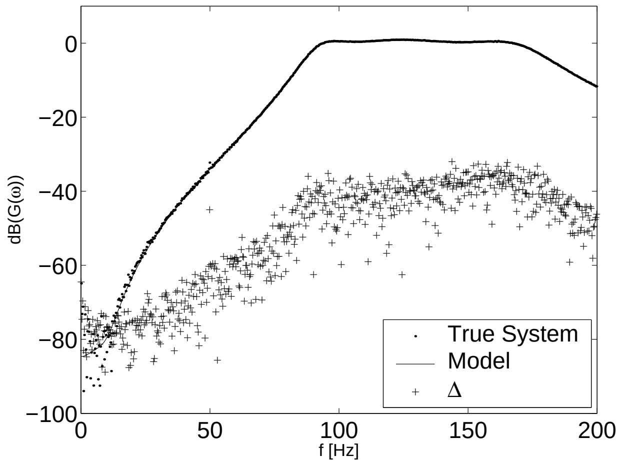

Figure 9: Identification result for the linear system inside the measured Hammerstein-Wiener system : comparison of the amplitude of the model and of the true system

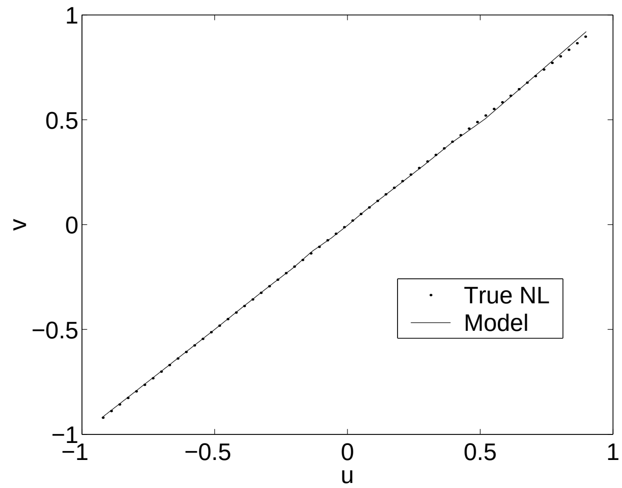

Figure 11: Identification result for the input nonlinearity $f(u)$ of the measured Wiener system

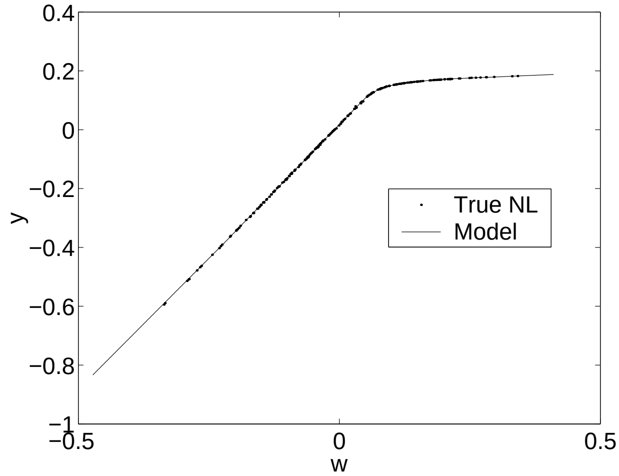

Figure 10: Identification result for the output nonlinearity $g(w)$ of the measured Hammerstein-Wiener system

Hammerstein-Wiener system where $v = f(u) = u$. This is verified in Figure 11 where the estimation neatly follows the line $v = u$ although the number of degrees of freedom for the estimator was quite high : 15 breakpoints were provided. Nonetheless, the algorithm didn't use these degrees of freedom to reduce the cost function by introducing an inexistent nonlinearity in its model. Hence, the Wiener structure has been revealed.

6 Conclusion

An algorithm to identify a Hammerstein-Wiener system has been presented. This algorithm generates starting values that can be used by a nonlinear minimization procedure that optimizes all parameters of the model at the same time. The method has been tested and demonstrated on measurement results.

If the true model structure is less complicated than the model structure used for identification, the identification behaves well : a Wiener system identified by this procedure yields a model that has a Wiener structure.

Acknowledgments

Philippe Crama is Aspirant of the Fund for Scientific Research Flanders (FWO, Fonds voor Wetenschappelijk Onderzoek Vlaanderen).

This work was supported by FWO-Vlaanderen, by the Flemish community under concerted action

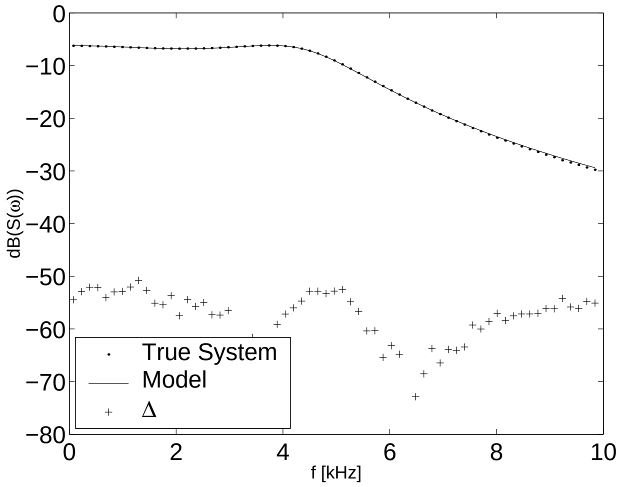

Figure 12: Identification result for the linear system inside the measured Wiener system : comparison of the amplitude of the model and of the true system

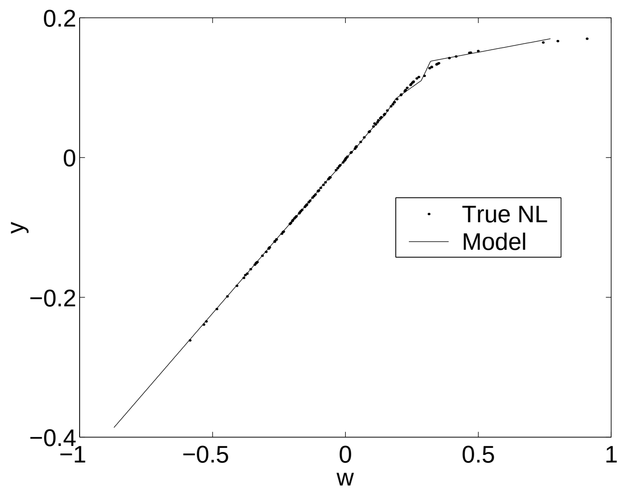

Figure 13: Identification result for the output nonlinearity $g(w)$ of the measured Wiener system

ILiNoS and by the Belgian government under IUAP-4/2.

References

[1] A. Hagenblad, *Aspects of the identification of Wiener models*, thesis no. 793, Linköping Stud. Sci. Technol., Linköping, Sweden, 1999

[2] Z. Q. Lang, *A nonparametric polynomial identification algorithm for the Hammerstein system*, IEEE Transactions on Automatic Control, Vol. 42, No. 10, 1997, pp. 1435–1441

[3] W. Greblicki, *Nonparametric approach to Wiener system identification*, IEEE Transactions on Circuits and Systems I-Fundamental Theory and Applications, Vol. 44, No. 6, 1997, pp. 538–545

[4] S. A. Billings, S. Y. Fakhouri, *Nonlinear system identification using the Hammerstein model*, International Journal of Systems Science, Vol. 10, No. 5, 1979, pp. 567–578

[5] S. A. Billings, S. Y. Fakhouri, *Identification of systems containing linear dynamic and static nonlinear elements*, Automatica, Vol. 18, No. 1, 1982, pp. 15–26

[6] S. A. Billings, K. M. Tsang, *Spectral analysis of block structured nonlinear systems*, Mechanical Systems Signal Processing, Vol. 4, No. 2, 1990, pp. 117–130

[7] E. W. Bai, *An Optimal Two-Stage Identification Algorithm for Hammerstein-Wiener Nonlinear Systems*, Automatica, Vol. 34, No. 3, 1998, pp. 333–338

[8] E. W. Bai, *A Blind Approach to the Hammerstein-Wiener Model Identification*, Automatica, Vol. 38, No. 6, 2002, pp. 967–979

[9] D. Bauer, B. Ninness, *Asymptotic properties of least-squares estimates of Hammerstein-Wiener models*, International Journal of Control, Vol. 75, No. 1, 2002, pp. 34–51

[10] HHJ Bloemen, TJJ van den Boom, HB Verbruggen, *Model-based predictive control for Hammerstein-Wiener systems*, International Journal of Control, Vol. 74, No. 5, 2001, pp. 482–495

[11] J. Schoukens, T. Dobrowiecki, R. Pintelon, *Parametric and Nonparametric Identification of Linear Systems in the Presence of Nonlinear Distortions — A Frequency Domain Approach*, IEEE Transactions on Automatic Control, Vol. 43, No. 2, 1998, pp. 176–190

[12] P. Guillaume, J. Schoukens, R. Pintelon, I. Kollár, *Crest-Factor Minimization Using Nonlinear Chebyshev Approximation Methods*, IEEE Transactions on Instrumentation and Measurement, Vol. 40, No. 6, 1991, pp. 982–989

[13] E. Van der Ouderaa, J. Schoukens, J. Renneboog, *Peak Factor Minimization of Input and Output Signals of Linear Systems*, IEEE Transactions on Instrumentation and Measurement, Vol. 37, No. 2, 1988, pp. 207–212

[14] R. Pintelon, J. Schoukens, *System Identification — A Frequency Domain Approach*, IEEE Press, New York, 2001

[15] T. P. Dobrowiecki, J. Schoukens, *Practical Choices in the FRF Measurement in Presence of Nonlinear Distortions*, IEEE Transactions on Instrumentation and Measurement, Vol. 50, No. 1, 2001, pp. 2–7

[16] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, Wiley, New York, 1980

[17] P. Crama, J. Schoukens, *Initial Estimates of Wiener and Hammerstein Systems Using Multisine Excitation*, IEEE Transactions on Instrumentation and Measurement, Vol. 50, No. 6, 2001, pp. 1791–1795